A GRAPH LAPLACIAN REGULARIZATION FOR HYPERSPECTRAL DATA UNMIXING

Rita Ammanouil, André Ferrari, Cédric Richard

Université de Nice Sophia-Antipolis, CNRS, Observatoire de la Côte d'Azur, France

ABSTRACT

This paper introduces a graph Laplacian regularization in the hyperspectral unmixing formulation. The proposed regularization relies upon the construction of a graph representation of the hyperspectral image. Each node in the graph represents a pixel's spectrum, and edges connect similar pixels. The proposed graph framework promotes smoothness in the estimated abundance maps and *collaborative* estimation between homogeneous areas of the image. The resulting convex optimization problem is solved using the Alternating Direction Method of Multipliers (ADMM). A special attention is given to the computational complexity of the algorithm, and Graph-cut methods are proposed in order to reduce the computational burden. Finally, simulations conducted on synthetic and real data illustrate the effectiveness of the graph Laplacian regularization with respect to other classical regularizations for hyperspectral unmixing.

Index Terms— Hyperspectral imaging, unmixing, graph Laplacian regularization, ADMM, sparse regularization.

1. INTRODUCTION

Hyperspectral sensors provide both a spatial and a spectral representation of a scene. They acquire images throughout the visible and Infrared portions of the spectrum, with a spectral resolution as narrow as 1 nm. Depending on the working distance of the hyperspectral camera, the spatial resolution can be of a few micrometers (laboratory measurements) up to a few meters (airborne remote sensing). As a result every pixel in the hyperspectral image is a vector of reflectance values also known as the pixel's spectrum. Unmixing [1] is one of the most prominent tools to analyze hyperspectral data. It consists of identifying the pure components in the captured scene, the so-called endmembers, and then estimating their spatial distributions, also known as their abundance maps. While there are also works in the literature that consider nonlinear mixing models [2, 3], most unmixing methods focus on the Linear Mixing model [4], where each pixel is modeled by a convex combination of the endmembers weighted by their abundances.

The purpose of this paper is to introduce the graph Laplacian regularization in the hyperspectral unmixing formulation. This is motivated by the intuition that pixels with similar spectral structure and similar spatial contextual information will have broadly similar abundances. Representing these pairwise similarities by edges gives rise to a graph, where each node represents a pixel. The resulting graph structure provides additional relational information which can improve the abundance estimation accuracy and complement existing *pixel-by-pixel* unmixing techniques. As we shall see further ahead in Section 3, the graph Laplacian regularization provides an elegant and flexible way to incorporate this information in the unmixing problem. This regularization has been widely used in many fields especially in semi-supervised learning also known as transductive learning. The potential of this regularization has been demonstrated for many applications including digit recognition and text classification [5], web-page categorization [6], hyperspectral data classification [7], manifold learning [8], and image denoising [9] to cite a few.

The proposed strategy is closely related to the work in [10] where the authors use a Total Variation (TV) regularization on top of sparse ℓ_1-norm regularized unmixing. Similarly to [10], this communication advocates the use of the graph Laplacian regularization on top of $\ell_{2,1}$-norm regularized unmixing. TV is restricted to the assumption of local spatial similarity, and assumes that a pixel is only similar to its four neighbors. However, the graph Laplacian regularization is more flexible in the sense that it allows to connect a pixel with as many other pixels in the image as long as they are similar. [11] extends the TV spatial regularization to nonlinear unmixing models. Several methods in the literature incorporate other spatial or spectral-spatial information in the unmixing problem such as [12–15]. For a detailed review of spectral unmixing methods and endmember extraction techniques with spatial information, the reader is referred to [16]. Very recently, the authors of [17] used the graph Laplacian regularization on top of sparse $\ell_{1/2}$ nonnegative matrix factorization (NMF) for blind unmixing. The algorithm uses alternate minimization in order to simultaneously estimate the endmembers and the abundances. In this work, we use the ADMM algorithm [18] which allows to take into account the abundances sum-to-one and positivity constraints, and a Group lasso regularizer frequently incorporated in unmixing to allow the the selection of endmembers from a large library [19]. Moreover, we exploit the graph structure and use an algorithm [20] similar to normalized graph-cuts [21] in order to partition the graph into several sub-graphs. Unmixing is then performed on each subgraph separately which allows to reduce the computational complexity of the algorithm.

The paper is organized as follows. Section 2 introduces the hyperspectral and graph frameworks, Section 3 incorporates the graph Laplacian regularization on top of sparse unmixing, Section 4 is devoted to testing the proposed approach using synthetic and real data. Finally, Section 5 concludes the paper.

2. HYPERSPECTRAL IMAGE TO GRAPH MAPPING

Let us first introduce the linear mixing model and some notations specific to the hyperspectral unmixing framework. In matrix form, the linear mixing model is given by

$$\boldsymbol{S} = \boldsymbol{RA} + \boldsymbol{E} \tag{1}$$

with $\boldsymbol{S} = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N)$, $\boldsymbol{R} = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_M)$, $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M)^\top$. Here, $\boldsymbol{s}_j$ is the L-dimensional spectrum of the j-th pixel, L is the number of frequency bands, $\boldsymbol{r}_i$ is L-dimensional spectrum of the

i-th endmember, M denotes the number of endmembers, $\boldsymbol{a}_i$ is the M-dimensional abundance map of the i-th endmember, N is the number of pixels in the image, and $\boldsymbol{E}$ is an additive Gaussian noise. All vectors are column vectors. Model (1) means that the (i, j)-th entry A_{ij} of $\boldsymbol{A}$ is the abundance of the endmember $\boldsymbol{r}_i$ in pixel $\boldsymbol{s}_j$. Two physical constraints on the abundances are usually considered, the non-negativity and sum-to-one constraints: $A_{ij} \geq 0$ for all (i, j), and $\sum_{i=1}^{M} A_{ij} = 1$ for all j.

The first step in the proposed graph-based unmixing approach consists in mapping the hyperspectral image to a graph $\boldsymbol{G}$ where each node represents a pixel's spectra. Let $\boldsymbol{W}$ be the $N \times N$ affinity matrix of the graph, the entries W_{ij} of $\boldsymbol{W}$ satisfy the following conditions. If pixels i and j are similar then $W_{i,j}$ is set to some positive value proportional to their degree of similarity. If pixels i and j are dissimilar then W_{ij} tends to zero. There are different heuristics for choosing W_{ij}. For example, this can be done by using a Gaussian kernel

$$W_{ij} = \exp\left(-\frac{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}{2\sigma^2} \right) \tag{2}$$

where σ is the kernel's bandwidth [22, 23]. In addition to the pixel's spectrum, each pixel can be defined by a vector of spatial features, for instance, the average of its surrounding area, its coordinates in the image. This spatial information leads to a second spatial affinity matrix which can be easily combined with the spectral one [7]. Finally, k-nearest neighbors and thresholding are commonly used in order to set to zero small weights in $\boldsymbol{W}$ [24]. The authors of [7, 22, 23] propose different strategies for defining an affinity matrix that takes into account both the spatial and the spectral information of a pixel.

3. LAPLACIAN REGULARIZED UNMIXING

As previously mentioned, we consider the following interpretation of the graph. If two nodes are connected, then they are likely to have similar abundances. We shall now incorporate this information in the unmixing problem using the graph Laplacian regularization. This leads to the following convex optimization problem:

$$\min_{\boldsymbol{A}} \quad \frac{1}{2}\|\boldsymbol{S} - \boldsymbol{R}\boldsymbol{A}\|_F^2 + \lambda \text{tr}(\boldsymbol{A}\mathcal{L}\boldsymbol{A}^\top) + \mu \sum_{k=1}^{M} \|\boldsymbol{a}_k\|_2$$
$$\text{subject to} \quad A_{ij} \geq 0 \quad \forall i, j$$
$$\sum_{i=1}^{N} A_{ij} = 1 \quad \forall j. \tag{3}$$

where $\mathcal{L}$ is the graph Laplacian matrix given by $\mathcal{L} = \boldsymbol{D} - \boldsymbol{W}$, $\boldsymbol{D}$ is a diagonal matrix with $\boldsymbol{D}_{ii} = \sum_{j=1}^{N} W_{ij}$, $\mu \geq 0$ and $\lambda \geq 0$ are two regularization parameters. The first term in (3) is a data fidelity term based on the ℓ_2-norm. The second term is the graph Laplacian regularization. To see the relevance of this regularization in (3), we rewrite it as follows [25]:

$$\text{tr}(\boldsymbol{A}\mathcal{L}\boldsymbol{A}^\top) = \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k \sim j} W_{jk}(A_{ij} - A_{ik})^2 \tag{4}$$

where $k \sim j$ indicates that pixels j and k are similar ($W_{jk} \neq 0$). For every abundance map (row in $\boldsymbol{A}$), this term penalizes the square of the difference between the abundances of similar pixels proportionally to their degree of similarity. This quantity can also be seen as a measure of the discrepancies between the abundance estimates weighted by their degree of similarity W_{jk}. The regularization parameter λ controls the extent at which similar pixels estimate similar abundances. The third term is the $\ell_{2,1}$-norm regularization also known as the Group lasso. We consider that $\boldsymbol{R}$ is a large dictionary of endmembers, and only few of these endmembers are present in the image. For this reason, we use the Group Lasso regularization to induce group sparsity [26] in the estimated abundance matrix by possibly driving several rows $\boldsymbol{a}_k$ of $\boldsymbol{A}$ to zero, as proposed in [19].

It is important to note that the first and second term of the cost function (3) are separable with respect to the columns and rows of $\boldsymbol{A}$ respectively. These two terms can be grouped into a single quadratic form. However, the resulting Quadratic Problem has $N \times M$ variables, and is not separable with respect to the columns or the rows of $\boldsymbol{A}$. The solution of problem (3) can be obtained in a simple and flexible manner using the Alternating Direction Method of Multipliers [18].

3.1. ADMM algorithm

We consider the canonical form and the following variable splitting:

$$\min_{\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}} \quad \frac{1}{2}\|\boldsymbol{S} - \boldsymbol{R}\boldsymbol{X}\|_F^2 + \lambda \text{tr}(\boldsymbol{Y}\mathcal{L}\boldsymbol{Y}^\top) + \mu \sum_{k=1}^{M} \|\boldsymbol{z}_k\|_2$$
$$+ \mathcal{I}(\boldsymbol{Z})$$
$$\text{subject to} \quad \boldsymbol{B}\boldsymbol{X} + \boldsymbol{C}\boldsymbol{Z} = \boldsymbol{F}$$
$$\boldsymbol{X} = \boldsymbol{Y} \tag{5}$$

with

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{I} \\ \boldsymbol{1}^\top \end{pmatrix}, \, \boldsymbol{C} = \begin{pmatrix} -\boldsymbol{I} \\ \boldsymbol{0}^\top \end{pmatrix}, \, \boldsymbol{F} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{1}^\top \end{pmatrix},$$

where $\mathcal{I}$ is the indicator of the positive orthant guarantying the positivity constraint, that is, $\mathcal{I}(\boldsymbol{Z}) = 0$ if $\boldsymbol{Z} \succeq \boldsymbol{0}$ and $+\infty$ otherwise. The constraints impose the consensus $\boldsymbol{X} = \boldsymbol{Y}$, $\boldsymbol{X} = \boldsymbol{Z}$, and the sum-to-one. In matrix form, the augmented Lagrangian for problem (5) is given by

$$\boldsymbol{L}_\rho(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{V}, \boldsymbol{\Lambda}) = \frac{1}{2}\|\boldsymbol{S} - \boldsymbol{R}\boldsymbol{X}\|_F^2 + \mu \sum_{k=1}^{M} \|\boldsymbol{z}_k\|_2 + \mathcal{I}(\boldsymbol{Z})$$
$$+ \lambda \text{tr}(\boldsymbol{Y}\mathcal{L}\boldsymbol{Y}^\top) + \text{tr}(\boldsymbol{V}^\top(\boldsymbol{X} - \boldsymbol{Y})) + \frac{\rho}{2}\|\boldsymbol{X} - \boldsymbol{Y}\|_F^2$$
$$+ \frac{\rho}{2}\|\boldsymbol{B}\boldsymbol{X} + \boldsymbol{C}\boldsymbol{Z} - \boldsymbol{F}\|_F^2 + \text{tr}(\boldsymbol{\Lambda}^\top(\boldsymbol{B}\boldsymbol{X} + \boldsymbol{C}\boldsymbol{Z} - \boldsymbol{F})) \tag{6}$$

where $\boldsymbol{\Lambda}$ and $\boldsymbol{V}$ are the matrices of the Lagrange multipliers, and ρ is the penalty parameter. The flexibility of the ADMM lies in the fact that it splits the initial optimization problem into three subproblems. At iteration $k + 1$, the ADMM algorithm is outlined by four sequential steps.

$\boldsymbol{X}$ **minimization step**: The augmented Lagrangian is quadratic with respect to $\boldsymbol{X}$. The minimizer has an analytical expression that is obtained by setting the gradient of the augmented Lagrangian with respect to $\boldsymbol{X}$ to zero:

$$\boldsymbol{X}^{k+1} = (\boldsymbol{R}^\top \boldsymbol{R} + \rho \boldsymbol{B}^\top \boldsymbol{B} + \rho \boldsymbol{I}_M)^{-1}(\boldsymbol{R}^\top \boldsymbol{S}$$
$$- \boldsymbol{B}^\top[\boldsymbol{\Lambda}^k + \rho(\boldsymbol{C}\boldsymbol{Z}^k - \boldsymbol{F})] - \boldsymbol{V}^k + \rho \boldsymbol{Y}^k). \tag{7}$$

$\boldsymbol{Y}$ **minimization step**: Similarly to the first step, $\boldsymbol{Y}^{k+1}$ is obtained by setting the gradient of the augmented Lagrangian with respect to $\boldsymbol{Y}$ to zero, which yields:

$$\boldsymbol{Y}^{k+1} = (\boldsymbol{V}^k + \rho \boldsymbol{X}^{k+1})(2\lambda \mathcal{L} + \rho \boldsymbol{I}_N)^{-1}. \tag{8}$$

Assume that we did not use $\boldsymbol{Y}$, and assigned the same ADMM variable $\boldsymbol{X}$ for both the fidelity term and the graph Laplacian regularization. In this case, the $\boldsymbol{X}$ minimization reduces to solving a Sylvester

equation [27]. The exact solution of this problem can not be computed efficiently due to the high dimensionality of the problem. In fact it requires the inversion of a $NM \times NM$ matrix where N and M can be both very large. Iterative methods have been proposed to perform this task [28]. These iterative methods are similar to the first two steps of our ADMM solution in the sense that the initial variable is split into two variables and alternating updates of these variables are performed.

Z minimization step: After discarding the terms that are independent of Z, the minimization of the augmented Lagrangian with respect to Z reduces to solving the following problem:

$$\min_{Z} \quad \mu \sum_{k=1}^{M} \|z_k\|_2 + \text{tr}(\Lambda^\top CZ) \\ + \frac{\rho}{2}\|BX + CZ - F\|_F^2 \tag{9}$$
$$\text{subject to} \quad Z \succeq 0.$$

This minimization step can be split into M problems given the structure of matrices B and C, one for each row of Z, that is,

$$\min_{z} \quad \frac{1}{2}\|z - v\|_2^2 + \alpha\|z\|_2 + \mathcal{I}(z) \tag{10}$$

where $v = x + \rho^{-1}\lambda$, $\alpha = \rho^{-1}\mu$. Vectors λ, x and z correspond to a given row of Λ, X and Z, respectively. The minimization problem (10) admits a unique solution given by the proximity operator of function $f(z) = \alpha\|z\|_2 + \mathcal{I}(z)$:

$$\begin{cases} z^* = 0 & \text{if } \|(v)_+\|_2 < \alpha \\ z^* = \left(1 - \frac{\alpha}{\|(v)_+\|_2}\right)(v)_+ & \text{otherwise,} \end{cases} \tag{11}$$

where $(\cdot)_+ = \max(0, \cdot)$. Operator (11) was recently used in [29, 30]. The derivation of this operator can be found in [30].

Update of the Lagrange multipliers Λ and V: The last step consists of updating the Lagrange multipliers Λ and V using the following expressions

$$\Lambda^{k+1} = \Lambda^k + \rho(BX^{k+1} + CZ^{k+1} - F), \\ V^{k+1} = V^k + \rho(X^{k+1} - Y^{k+1}). \tag{12}$$

As suggested in [18], a reasonable stopping criterion for this iterative algorithm is that the primal and dual residuals must be smaller than some tolerance thresholds.

3.2. A note on complexity

Finally, we pay particular attention to the computational complexity of the resulting ADMM algorithm. The most expensive step is the Y minimization, since it requires solving a linear system with N variables (8), N being very large in real images. We propose to exploit the graph representation of the pixels and apply the algorithm of [20] in order to partition the nodes of the graph into k clusters. This allows to approximately solve (8) by solving k smaller linear systems, where the number of variables is now smaller than N. Note that the overall optimization problem (3) can not be similarly separated into k distinct problems due to the group lasso regularization.

The purpose of this step is to reduce the computational complexity of (8) while preserving the global knowledge captured by the graph structure. For this reason, the segmentation must be conservative, in the sense that k should not be very large. In fact, a segmentation into too many clusters with strong connections between inter-cluster pixels could create undesirable cluster-like artifacts in the abundance maps.
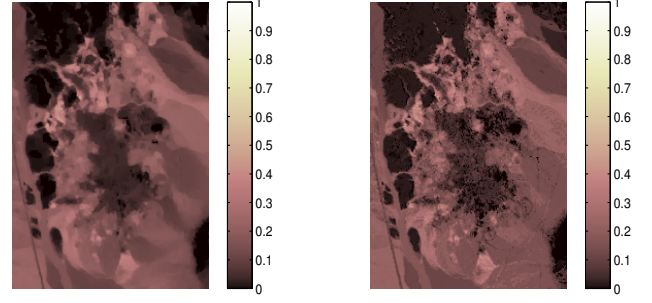
Fig. 2. Abundance maps for Alunite obtained using SUnSAL-TV and the proposed GLUP-Lap approach respectively.

4. EXPERIMENTS

The performance of the proposed approach was first evaluated using two simulated data sets, namely, Data1 and Data2 designed with different levels of homogeneity. Data1 is the same data set used in the experiments of [10, 11]. The image consists of 75×75 pixels generated using 5 endmembers $[r_1, r_2, \cdots r_5]$. The library R used in the experiments contains in total 230 endmembers with 224 spectral bands extracted from the USGS site[1]. The background is a mixture of the 5 endmembers with the following abundances $[0.1149\, 0.0741\, 0.2003\, 0.2055\, 0.4051]^\top$. There are 25 squares in the image disposed in a 5×5 grid fashion (see Figure 1). Each square is an homogeneous surface where its pixels have the same abundances. Most of the abundances in Data1 verify the assumption of local consistency. Data2 is generated similarly to Data1, except that it is created using 15 distinct endmembers, and the squares in each row are identical. In addition to local consistency, there exists distant homogeneous surfaces in Data2 that are identical. As a result a pixel has local similar neighbors and distant ones too.
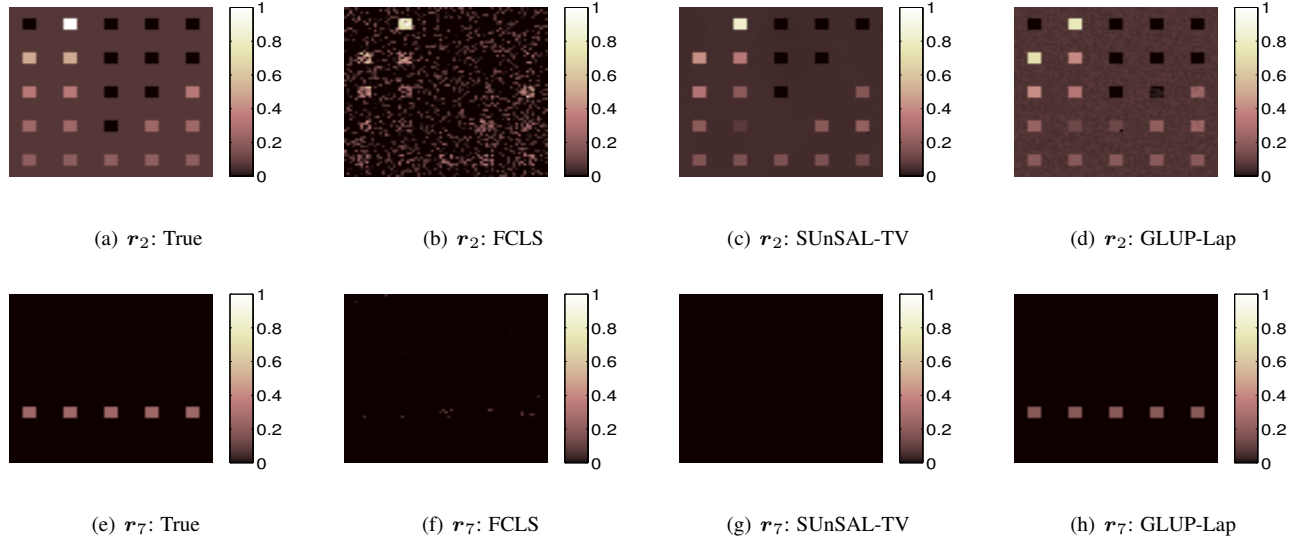
The first step in the proposed approach consists of defining the affinity matrix W. In all the experiments we simply threshold the square of the spectral distance and set the weights according to (13):

$$\begin{cases} W_{ij} = 1 & \text{if } \|s_i - s_j\|_2^2 < d_{\min}^2 \\ W_{ij} = 0 & \text{otherwise,} \end{cases} \tag{13}$$

where $d_{\min}^2$ represents the maximum squared spectral distance required in order to consider that two pixels are similar. As previously explained in Section 2, there are different heuristics for choosing the weights. (13) was sufficient in our experiments to demonstrate the effectiveness of the method. The algorithm described in [31] is then used to cut the graph into 10 disjoint subgraphs and then unmixing is performed on each subgraph.

We compared the performances of FCLS [4] and SUnSAL-TV [10] with the proposed approach denoted by GLUP-Lap (Group Lasso with Unit sum, Positivity constraints and graph Laplacian regularization). We used the Root Mean Square Error (RMSE) defined as $\text{RMSE} = (\frac{1}{NL} \times \|\hat{A} - A\|_F^2)^{\frac{1}{2}}$ as the evaluation metric. We tested SUnSAL-TV and GLUP-Lap for different combinations of the sparsity and the spatial tuning parameters μ and λ. Table 1 reports the best performance of each algorithm for a given data set and a given SNR with the corresponding optimal pair of regularization parameters. GLUP-Lap requires the tuning of an additional parameter d_{min}^2 which is also reported in the table. Both, SUnSAL-TV and GLUP-Lap, outperformed FCLS. GLUP-Lap had the lowest RMSE

---

[1]http://speclab.cr.usgs.gov/spectral-lib.html

| (a) r_2: True | (b) r_2: FCLS | (c) r_2: SUnSAL-TV | (d) r_2: GLUP-Lap |

| (e) r_7: True | (f) r_7: FCLS | (g) r_7: SUnSAL-TV | (h) r_7: GLUP-Lap |

Fig. 1. First row: Abundance maps for endmember 2 in Data1 obtained with SNR = 30 dB. Second row: Abundance maps for endmember 7 in Data2 obtained with SNR = 30 dB. From left to right: The true abundance map, FCLS, SUnSAL-TV, GLUP-Lap. The parameters are the reported in Table 1.

for all cases. As the SNR increases the rate at which GLUP-Lap improves with respect to FCLS increases. This is due to the fact that the observations contain less noise, thus the adjacency matrix becomes more reliable. The simulations performed with Data2 show that this data set is more difficult than the previous since it contains a large number of endmembers: 15 compared to 5 in Data1. As before, GLUP-Lap outperformed FCLS and SUnSAL-TV. It is important to note that GLUP-Lap and SUnSAL-TV were run under the same ADMM conditions. The penalty parameter was set to 0.05, and the maximum number of iterations to 200 in both algorithms. The first row of Figure 1 shows the true abundance map of the second endmember in Data1, and the estimated maps obtained with FCLS, SUnSAL-TV and GLUP-Lap with SNR=30dB. It can be seen from these maps that both SUnSAL-TV and GLUP-Lap estimated smooth abundance maps compared to FCLS, with SUnSAL-TV having the smoothest map. However, the squares that were not correctly estimated by SUnSAL-TV were better estimated with GLUP-Lap. This is possibly due to the fact that the pixels in these squares were encouraged by the graph to have similar estimates and thus appeared as consistent blocks in the GLUP-Lap result. The same observation can be made for the abundance map of r_7 in Data2. This endmember is only present in squares of the fourth row. The second row of Figure 1 shows that FCLS was not able to correctly estimate the abundance of this endmember in the figure. SUnSAL-TV, possibly due to the links it form with its surrounding estimates also failed to correctly estimate its abundances. Despite the difficulty of this abundance map, GLUP-Lap perfectly recovered the abundances. Even if the 5 squares are separated, their pixels are possibly connected in the graph and collaboratively estimate their abundances.

We also tested the proposed approach on a subset of the Cuprite scene [2] provided by the AVIRIS spectrometer. This scene was captured over a mining district in Nevada, the subset we use has 191×250 pixels and 188 spectral bands over the wavelength interval $400-2500$ nm. Figure 2 shows the abundance maps of Alunite obtained

[2] available at http://www.ehu.eus/ccwintco

using SUnSAL-TV and GLUP-Lap respectively. In the former case, μ and η were both set to 10^{-3} and the execution time was 1281 seconds. In the latter case, μ and η were set to 10^{-3} and 0.1 respectively, d^2_{min} was set to 0.005, and the execution time was 5275 seconds. It can be seen from Figure 2 that the abundance maps estimated by GLUP-Lap are less smooth than SUnSAL-TV. However, they conserved relatively more details.

Table 1. RMSE obtained with different values of the SNR, with the optimal values of the couple $(\mu; \lambda)$ for SUnSAL-TV and GLUP-Lap, the penalty parameter was set to $\rho = 0.05$ for both algorithms.

	SNR 20 dB	SNR 30 dB	SNR 40 dB
	Data1		
FCLS	0.0262	0.0173	0.0101
SUnSAL TV	0.0156 $(0.05; 0.05)$	0.0075 $(5\ 10^{-3}; 0.01)$	0.0034 $(10^{-3}; 5\ 10^{-3})$
GLUP Lap	0.0152 $(0.01; 0.5)$ $d^2_{min} = 2.5$	0.0049 $(5\ 10^{-4}; 0.5)$ $d^2_{min} = 0.3$	0.0012 $(5\ 10^{-5}; 0.5)$ $d^2_{min} = 0.05$
	Data2		
FCLS	0.0307	0.0240	0.0151
SUnSAL TV	0.0250 $(0.05; 0.3)$	0.0132 $(10^{-4}; 0.005)$	0.0073 $(5\ 10^{-5}; 10^{-3})$
GLUP Lap	0.0174 $(0.01; 1)$ $d^2_{min} = 1.8$	0.0078 $(10^{-4}; 1)$ $d^2_{min} = 0.5$	0.0023 $(5\ 10^{-5}; 1)$ $d^2_{min} = 0.5$

5. CONCLUSION

In this work we used an affinity matrix of the image in order to incorporate the graph Laplacian regularization within the sparse unmixing formulation. We showed that the resulting graph regularized framework has potential in improving the abundances' estimation accuracy and creates more consistent areas at the local and global level. Future work includes studying the potential of spatial-spectral weights for further improving the performance.

6. REFERENCES

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.

[2] N. Dobigeon, J.-Y. Tourneret, C. Richard, J.-C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 82–94, 2014.

[3] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 480–492, 2013.

[4] D. C. Heinz and C. I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hypersectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.

[5] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schlkopf, "Learning with local and global consistency," *NIPS*, vol. 16, no. 16, pp. 321–328, 2004.

[6] T. Zhang, A. Popescul, and B. Dom, "Linear prediction models with graph regularization for web-page categorization," in *Proc. ACM SIGKDD*, 2006, pp. 821–826.

[7] G. Camps-Valls, T. B. Marsheva, and D. Zhou, "Semi-supervised graph-based hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3044–3054, 2007.

[8] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," *The Journal of Machine Learning Research*, vol. 7, pp. 2399–2434, 2006.

[9] A. Kovac and A. Smith, "Nonparametric regression on a graph," *Journal of Computational and Graphical Statistics*, vol. 20, no. 2, pp. 432–447, 2011.

[10] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4484–4502, 2012.

[11] J. Chen, C. Richard, and P. Honeine, "Nonlinear estimation of material abundances in hyperspectral images with l1-norm spatial regularization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2654 – 2665, 2014.

[12] S. Jia and Y. Qian, "Spectral and spatial complexity-based hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 3867–3879, 2007.

[13] A. Castrodad, Z. Xing, J. Greer, E. Bosch, L. Carin, and G. Sapiro, "Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4263–4281, 2011.

[14] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4239–4247, 2011.

[15] A. Zare, "Spatial-spectral unmixing using fuzzy local information," in *Proc. IEEE IGARSS*, 2011, pp. 1139–1142.

[16] C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," *Remote Sensing of Environment*, vol. 149, pp. 70–87, 2014.

[17] L. Tong, J. Zhou, X. Bai, and Y. Gao, "Dual graph regularized nmf for hyperspectral unmixing," in *Proc. DICTA*, Nov. 2014.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[19] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 341–354, 2013.

[20] A. Ng, M. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," *Advances in neural information processing systems*, vol. 2, pp. 849–856, 2002.

[21] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[22] D. Gillis and J. Bowles, "Hyperspectral image segmentation using spatial-spectral graphs," in *Proc. SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2012, pp. 83901Q–83901Q.

[23] Y. Zhang, H. L. Yang, D. Lunga, S. Prasad, and M. Crawford, "Spatial context driven manifold learning for hyperpsectral image classification," in *Proc. IEEE WHISPERS*, 2014.

[24] A. Argyriou, M. Herbster, and M. Pontil, "Combining graph laplacians for semi–supervised learning," *NIPS*, 2005.

[25] B. Mohar and Y. Alavi, "The laplacian spectrum of graphs," *Graph theory, combinatorics, and applications*, vol. 2, pp. 871–898, 1991.

[26] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical methodology)*, vol. 68, no. 1, pp. 49–67, February 2006.

[27] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation $AX + XB = C$," *Communications of the ACM*, vol. 15, no. 9, pp. 820–826, 1972.

[28] F. Ding and T. Chen, "Gradient based iterative algorithms for solving a class of matrix equations," *IEEE Transactions on Automatic Control*, vol. 50, no. 8, pp. 1216–1221, 2005.

[29] E. Thibaut, F. Soulez, and L. Denis, "Exploiting spatial sparsity for multiwavelength imaging in optical interferometry," *Journal of the Optical Society of America A*, vol. 30, no. 2, pp. 160–170, 2013.

[30] R. Ammanouil, A. Ferrari, C. Richard, and D. Mary, "Blind and fully constrained unmixing of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5510–5518, 2014.

[31] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.